## REMARKS

By this amendment, claims 1-8, 10, 12-25, 27, and 29-34 are pending, in which claims 9, 11, 26, and 28 are cancelled without prejudice or disclaimer and claims 1, 10, 18, and 27 are amended. Entry after final is proper since the amendment merely incorporates dependent claims 9, 11, 26, and 28 into respective independent claims 1, 10, 18, and 27, thereby avoiding the introduction of new matter or requiring a new search or further consideration.

The Office Action mailed March 25, 2003, rejected claims 1-5, 9-13, 16-22, 26-30, and 33-24 obvious under 35 U.S.C. § 103 based on *Hall et al.* (Hall et al., "Generating Fuzzy Rules from Data," *IEEE*, 1996), claims 6, 14, 23, 31 as obvious over *Hall et al.* in view of *Shafer et al.* (Shafer et al., "SPRINT: A Scalable Parallel Classifier for Data Mining," *Proceedings of the 22nd VLDB Conference*, 1996), and claims 7-8, 15, 24-25, and 32 as obvious over *Hall et al.* in view of *Choe et al.* (Choe et al., "On the Optimal Choice of Parameters in a Fuzzy C-Means Algorithm," *IEEE*, 1992). These rejections are respectfully for at least the following reasons.

### CLAIMS 1-8 AND 18-25

The rejection of claims 1-8 and 18-25 is respectfully traversed because *Hall et al.*, individually or in combination with *Shafer et al.* and *Choe et al.*, fail to teach or otherwise the limitations of the claims. For example, independent method claim 1 (whose limitations are mirrored in independent computer-readable medium claim 18) sets forth:

> 1. (Currently Amended) A method for refining a node of a decision tree associated with a plurality of data characterized by a plurality of features, comprising:
>
> selecting a feature from among the features characterizing the data associated with the node;
>
> performing a cluster analysis along the selected feature to group the data into one or more clusters;
>
> constructing one or more arcs of the decision tree at the node respectively for each of the one or more clusters; and

projecting the data in each of the clusters, wherein the projected data are characterized by the plurality of the features but for the selected feature; and

**recursively** performing the steps of selecting a feature and **performing the cluster analysis on the projected data in each of the clusters**.

Accordingly, independent claims 1 and 18 recites a way of refining a node in a decision tree by selecting a feature of those that characterize the data associated with the node, then performing a cluster analysis along the selected feature, and then constructing arcs of the decision tree for each of the clusters. Furthermore, the claims recite "recursively ... performing the cluster analysis on the projected data in each of the clusters." Thus, another cluster analysis is recursively performed on projected data in each of the clusters, enabling the decision to be built "on the fly" (see Spec., p. 6).

By contrast, *Hall et al.* does not show this way of building a decision tree. Rather, *Hall et al.* is directed to a method of developing of fuzzy rules from continuous valued data by building a decision tree in accordance with the C4.5 algorithm (Abstract, p. 1757, col. 1). However, *Hall et al.* recognizes that the "C4.5 algorithm tree algorithm **requires crisp** class assignments for all objects. It is **necessary** to partition the continuous output values into a effect set of **discrete** output classes." (Section 2.1, p. 1758, col. 1, emphasis added). Accordingly, *Hall et al.* propose to preprocess the data first by applying a fuzzy c-means clustering to determine the discrete classes, and then feeding the discrete classes into the C4.5 algorithm: "After a discrete class has been created for each example, as discussed in Section 2.1, C4.5 may be used to create a decision tree." (Section 3, p. 1759, col. 1).

Accordingly, whatever cluster analysis that is performed in *Hall et al.* is performed before, not during, building the decision tree, and thus there is no teaching or suggestion "recursively ... performing the cluster analysis on the projected data in each of the clusters." The portions of *Hal et al.* cited in the Office Action (viz. FIG. 1 and pp. 1757-1758) merely show a

decision tree without any suggestion of the recursively performed cluster analysis The remaining references, *Shafer et al.* and *Choe et al.*, also fail to teach this aspect of claims 1-9 and 18-26.

CLAIMS 3, 10, 12-17, 20, 27, AND 29-34

*Hall et al.*, alone or in combination with *Shafer et al.* and *Choe et al.*, fail to teach or suggest the limitations of claims 3, 10, 12-17, 20, 27, and 29-34, which specify using a "partition coefficient" to select one of the features.

As explained above, *Hall et al.* discloses a method of generating fuzzy rules from data by first performing a fuzzy cluster analysis to determine crisp, discrete classes for the data and then applying the C4.5 decision tree algorithm to the discrete classes. Since the C4.5 decision tree algorithm requires discrete classes, the C4.5 algorithms selects its features to build the decision tree based on the "highest information gain associated with it" (Section 2, p. 1758, col. 1). Although *Hall et al.* fails to give any details about how to calculated an "information gain," which, by itself, is deficiently in establishing a factual basis for the recited "partition coefficient," the specification of the present application explains on p. 3, line 17, that the "information is calculated by finding the average entropy of each attribute." However, average entropy is not a partition coefficient.

The Office Action contention that "the maximum information gain is chosen among the calculated information gain ratio as the partition coefficient" does not support the rejection because an information gain is not the same thing as a partition coefficient–as those terms are understood by one of skill in the art. Well-settled case law holds that the words of a claim must be read as they would be interpreted by those of ordinary skill in the art. *In re Baker Hughes Inc.*, 215 F.3d 1297, 55 USPQ2d 1149 (Fed. Cir. 2000); *In re Morris*, 127 F.3d 1048, 1054, 44 USPQ2d 1023, 1027 (Fed. Cir. 1997); M.P.E.P. 2111.01. "Although the PTO must give claims

their broadest reasonable interpretation, this interpretation must be consistent with the one that those skilled in the art would reach." *In re Cortright,* 165 F.3d 1353, 1369, 49 USPQ2d 1464, 1465 (Fed. Cir. 1999). Therefore, the rejection cannot be salvaged by redefining a "partition coefficient" as information gain, when "partition coefficient" already has a definite meaning that is not an information gain.

The remaining references, *Shafer et al.* and *Choe et al.*, also fail to teach this aspect of claims 3, 10, 12-17, 20, 27, and 29-34 and were not cited for that purpose.

CLAIMS 7-8, 15, 24-25, AND 32

The dependent claims are allowable for at least the same reasons as their independent claims and are individually on their own merits. For example, dependent claims 7-8, 15, 24-25, and 32 cover the element of "calculating a domain **ratio** of a difference in domains limits of the data **over** a difference in domain limits of a superset of the data." Accordingly, the domain ratio is recited to be a ratio of one difference over another.

The Office Action recognized correctly that *Hall et al.* fails to disclose this element, but incorrectly relies on *Choe et al.* for this feature. Specifically, the Office Action cites step 6 of *Choe et al.*'s algorithm, which states: "return to Step 3 if $\| U^{(l+1)} - U^{(l)} \| > \varepsilon$."[1] However, this condition is a **difference** of two quantities, $U^{(l+1)}$ and $U^{(l)}$, not a ratio of two differences as recited in claims 7-8, 15, 24-25, and 32. Furthermore, the U in *Choe et al.* is a "fuzzy c-partition of X" not the domain limit in these claims.

In response, the Office Action asserts that "instead of the different between two consecution U, a domain ratio could be used and still give the same result." However, claims 7-8, 15, 24-25, and 32 specifically require a "domain ratio." No domain ratio is disclosed in *Choe*

*et al.* and the Office Action admits so much. In rejecting a claim under 35 U.S.C. § 103, the Examiner is required to provide a factual basis to support the obviousness conclusion. *In re Warner*, 154 USPQ 173 (CCPA 1967); *In re Lunsford*, 148 USPQ 721 (CCPA 1966); *In re Freed*, 165 USPQ 570 (CCPA 1970). The Examiner is required to show that all the claim limitations are taught or suggested by the references. *In re Royka*, 180 USPQ 580 (CCPA 1974); *In re Wilson*, 165 USPQ 494 (CCPA 1970). It is therefore clear that none of the references teach or suggest the recited "domain ratio," so there is no factual basis to support the obviousness rejection.

Therefore, the present application, as amended, overcomes the objections and rejections of record and is in condition for allowance. Favorable consideration is respectfully requested. If any unresolved issues remain, it is respectfully requested that the Examiner telephone the undersigned attorney at 703-425-8516 so that such issues may be resolved as expeditiously as possible.

Respectfully Submitted,

DITTHAVONG & CARLSON, P.C.

6/25/2003
Date

Stephen C. Carlson
Attorney/Agent for Applicant(s)
Reg. No. 39929

10507 Braddock Rd
Suite A
Fairfax, VA 22032
Tel. 703-425-8516
Fax. 703-425-8518

[1] Due to typographical limitations of the Applicant's word processor, the "not less than or equal to" symbol (≤ with a stroke through it) of *Choe et al.* is replaced by the equivalent "greater than" (>) symbol.

## APPENDIX

1. (Currently Amended) A method for refining a node of a decision tree associated with a plurality of data characterized by a plurality of features, comprising:

selecting a feature from among the features characterizing the data associated with the node;

performing a cluster analysis along the selected feature to group the data into one or more clusters; [and]

constructing one or more arcs of the decision tree at the node respectively for each of the one or more clusters;

projecting the data in each of the clusters, wherein the projected data are characterized by the plurality of the features but for the selected feature; and

recursively performing the steps of selecting a feature and performing the cluster analysis on the projected data in each of the clusters.

9. (Canceled) [The method according to claim 1, further comprising the steps of:

projecting the data in each of the clusters, wherein the projected data are characterized by the plurality of the features but for the selected feature; and

recursively performing the steps of selecting a feature and performing the cluster analysis on the projected data in each of the clusters.]

10. (Currently Amended) A method for generating a decision tree for a plurality of data characterized by a plurality of features, comprising:

performing a plurality of cluster analyses along each of the features to calculate a maximal cluster validity measure, said maximal cluster validity measure corresponding to one of the features;

selecting the one of the features corresponding to the maximal cluster validity measure;

subdividing the data into one or more groups based on the selected feature; and

building the decision tree based on the one or more groups,

wherein the step of performing the cluster analyses along each of the features to calculate a maximal cluster validity measure includes the performing the steps of:

for each of the features, performing a plurality of cluster analyses along said each of the features for a plurality of cluster numbers to calculate respective partition coefficients; and

determining the maximal cluster validity measure from among the partition coefficients.

11. (Canceled) [The method according to claim 10, wherein the step of performing the cluster analyses along each of the features to calculate a maximal cluster validity measure includes the performing the steps of:

for each of the features, performing a plurality of cluster analyses along said each of the features for a plurality of cluster numbers to calculate respective partition coefficients; and

determining the maximal cluster validity measure from among the partition coefficients.]

18. (Currently Amended) A computer-readable medium bearing instructions for refining a node of a decision tree associated with a plurality of data characterized by a plurality of features, said instructions being arranged to cause one or more processors upon execution thereby to perform the steps of:

selecting a feature from among the features characterizing the data associated with the node;

performing a cluster analysis along the selected feature to group the data into one or more clusters; [and]

constructing one or more arcs of the decision tree at the node respectively for each of the one or more clusters;

projecting the data in each of the clusters, wherein the projected data are characterized by the plurality of the features but for the selected feature; and

recursively performing the steps of selecting a feature and performing the cluster analysis on the projected data in each of the clusters.

26. (Canceled) [The computer-readable medium according to claim 18, wherein said are further arranged to the one or more processors upon execution thereby to perform the steps of:

projecting the data in each of the clusters, wherein the projected data are characterized by the plurality of the features but for the selected feature; and

recursively performing the steps of selecting a feature and performing the cluster analysis on the projected data in each of the clusters.]

27. (Currently Amended) A computer-readable medium bearing instructions for generating a decision tree for a plurality of data characterized by a plurality of features, said instructions being arranged to cause one or more processors upon execution thereby to perform the steps of:

performing a plurality of cluster analyses along each of the features to calculate a maximal cluster validity measure, said maximal cluster validity measure corresponding to one of the features;

selecting the one of the features corresponding to the maximal cluster validity measure;

subdividing the data into one or more groups based on the selected feature; and

building the decision tree based on the one or more groups,

wherein the step of performing the cluster analyses along each of the features to calculate a maximal cluster validity measure includes the performing the steps of:

for each of the features, performing a plurality of cluster analyses along said each of the features for a plurality of cluster numbers to calculate respective partition coefficients; and

determining the maximal cluster validity measure from among the partition coefficients.

28. (Canceled) [The computer-readable medium according to claim 27, wherein the step of performing the cluster analyses along each of the features to calculate a maximal cluster validity measure includes the performing the steps of:

for each of the features, performing a plurality of cluster analyses along said each of the features for a plurality of cluster numbers to calculate respective partition coefficients; and

determining the maximal cluster validity measure from among the partition coefficients.]